Universitetet i Sørøst-Norge

Is It Really Necessary to Go Beyond A Fairness Metric for Next-Generation Congestion Control?

Safiqul Islam[1], Kristian Hiorth[2], Carsten Griwodz[2], Michael Welzl[2]

[1]University of South-Eastern Norway

[2]University of Oslo

ANRW 2022
Philadelphia, PA

Introduction

- The evaluation of congestion control mechanisms encompasses an examination of fairness
  - Using Jains Fairness Index

$$\text{JFI} = \frac{(\sum_{i=1}^{N} x_i(t))2}{N\sum_{i=1}^{N} x_i(t)^2}$$

Introduction

- The evaluation of congestion control (CC) mechanisms encompasses an examination of fairness
  - Using Jains Fairness Index

$$\text{JFI} = \frac{(\sum_{i=1}^{N} x_i(t))2}{N \sum_{i=1}^{N} x_i(t)^2}$$

- To judge how multiple instances of the same CC mechanism interoperate

Introduction

- Also used to evaluate whether a new mechanism is fit for deployment
  – Evaluating fairness when a new mechanism competes with the prevalent CC mechanism
- Briscoe argues that fairness should be defined in relation to cost, per economic entity – not per flow
  – However, it is still common to evaluate mechanisms on the basis of flow-rate fairness

Briscoe, Bob. "Flow rate fairness: Dismantling a religion." *ACM SIGCOMM Computer Communication Review* 37.2 (2007): 63-74.

Does such a fairness test indeed provide a good reasoning about the deployment of a new congestion control mechanism?

Harm concept

- Ware et al. suggested to use the concept of how harmful a new entrant CC algorithm is to incumbent CC algorithms
  - Developers should also focus on various performances metrics
    - delay, loss and flow completion time
- The harm concept seems practical but its practical merit hasn't been demonstrated
  - Requires more experimental data than the calculation of JFI

Ware, Ranysha, et al. "Beyond Jain's Fairness Index: Setting the Bar For The Deployment of Congestion Control Algorithms." *Proceedings of the 18th ACM Workshop on Hot Topics in Networks*. 2019.

We provide the first evaluation of using a fairness metric vs. using harm with representative CC mechanisms.

Outline

- How to calculate harm
- Representation suitable for comparison
- Measurement setup
- Results

How to calculate harm

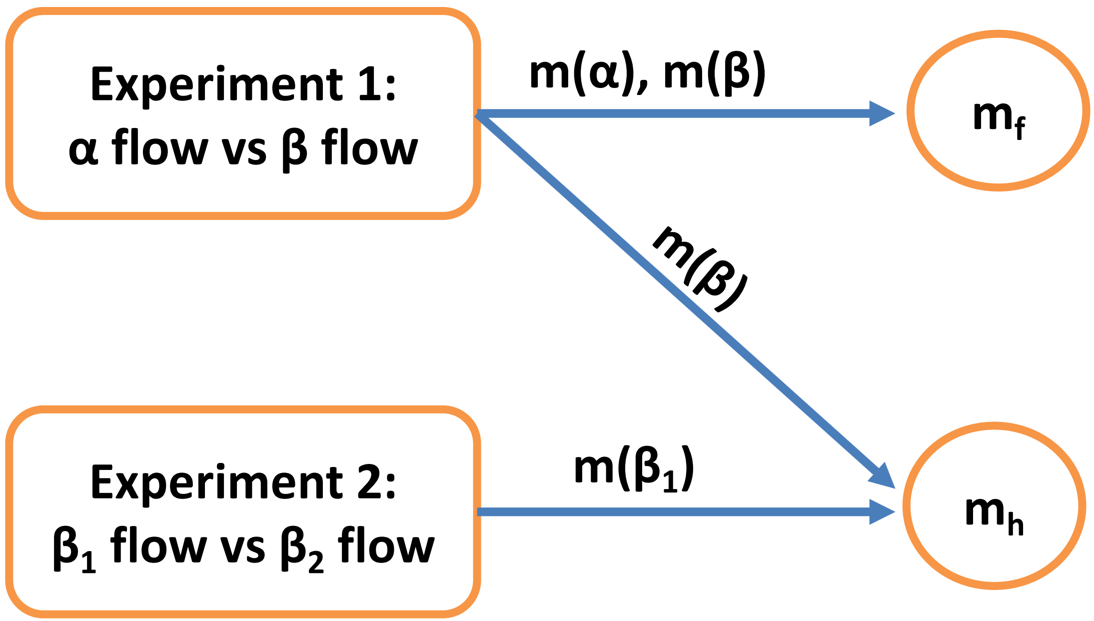
From the harm paper:

“We suggest that, if the harm done by a new CCA alpha to a widely-deployed CCA beta is comparable or less than the harm done when beta competes against beta, we should consider it acceptable to deploy.”

Representation suitable for comparison

- Based on this, we carry out two tests for all our scenarios
  - First: flow α (a new cc mechanism) competing with flow β (baseline cc)
  - Second: two baseline flows β_1 and β_2 competing with each other
- Mapping from a flow to a specific measurement is referred to as: m: flow --> metric value

Representation suitable for comparison



Negative values correspond to: α causing much harm to β
Zero: no harm
Positive values: β harms α when they compete

Measurement setup

- Ran experiments in our teacup physical testbed
- Varied link capacity to 10, 25, 50, 75 and 100 Mbps
- RTT varied to 10, 20, 50, and 100 ms
- Queue size: set to half a BDP and a full BDP for each bandwidth and delay case

Measurement setup

- Four different CC mechanisms were used based on their level of aggression and congestion signal they use:

Name	Aggression	Loss-based	Delay-based
Reno	+	x	o
Cubic	++	x	o
BBR	+++	x♣	x
Vegas	-	o†	x

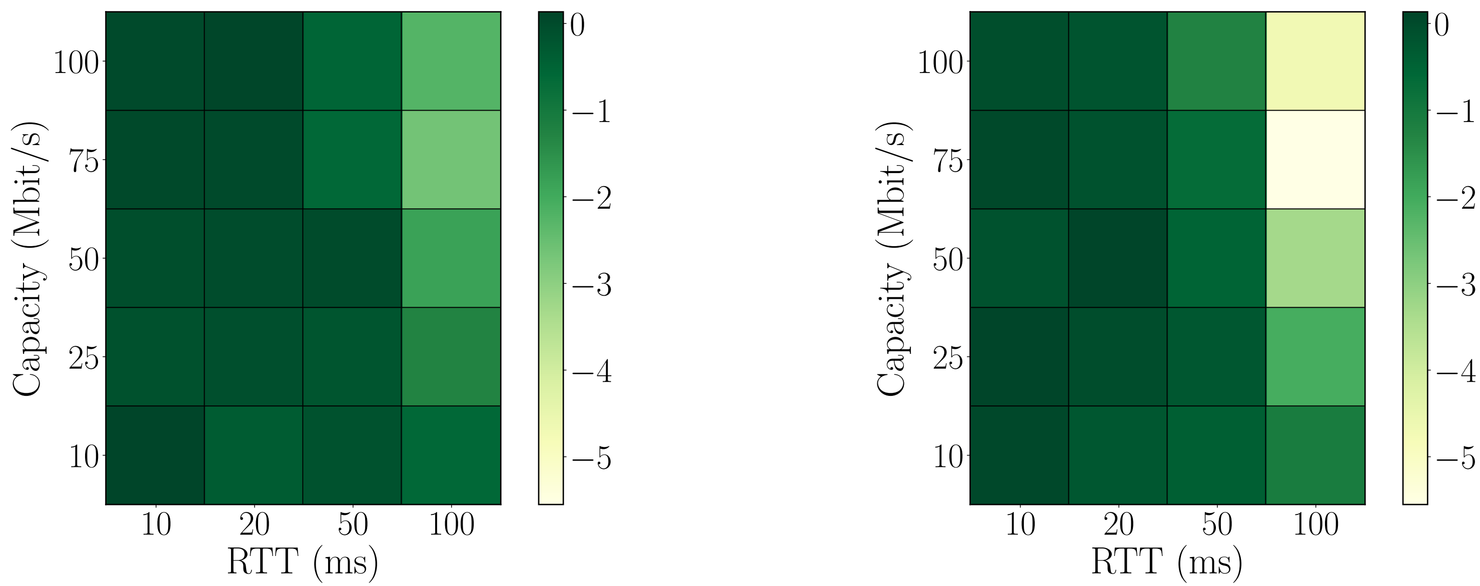
♣ while BBRv1 ignored explicit loss notifications – BBR has always implicitly used loss (in addition to delay)

† Vegas does not ignore loss but it is primarily delay-based

Results

- Investigate harm metric for Cubic vs Reno and Cubic vs BBR flows
- Investigate how the metric behaves when we run a loss-based flow against a delay-based flow
- Share our experience whether the deployability of new CC mechanism could be judged by a harm-based approach

Cubic vs Reno: identify and eliminate scenarios where Cubic falls back to linear TCP-like growth

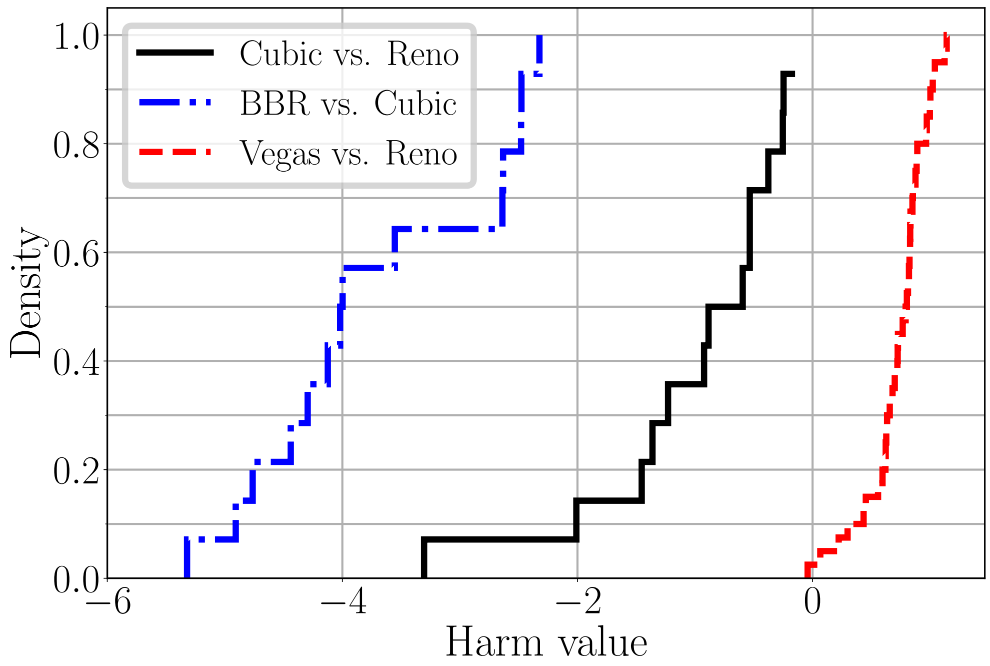


Queue sized to half a BDP

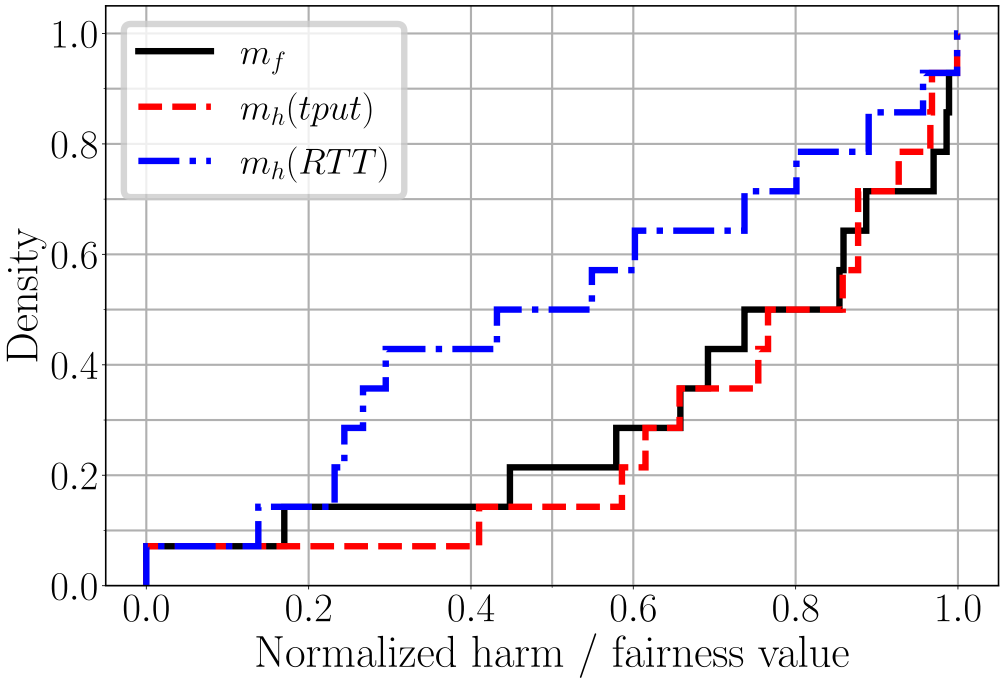
Queue sized to a full BDP

Fairness comparison between Cubic and Reno is interesting at RTT= 50 ms and capacity >= 75 Mpbs as well as tests where RTT=100ms

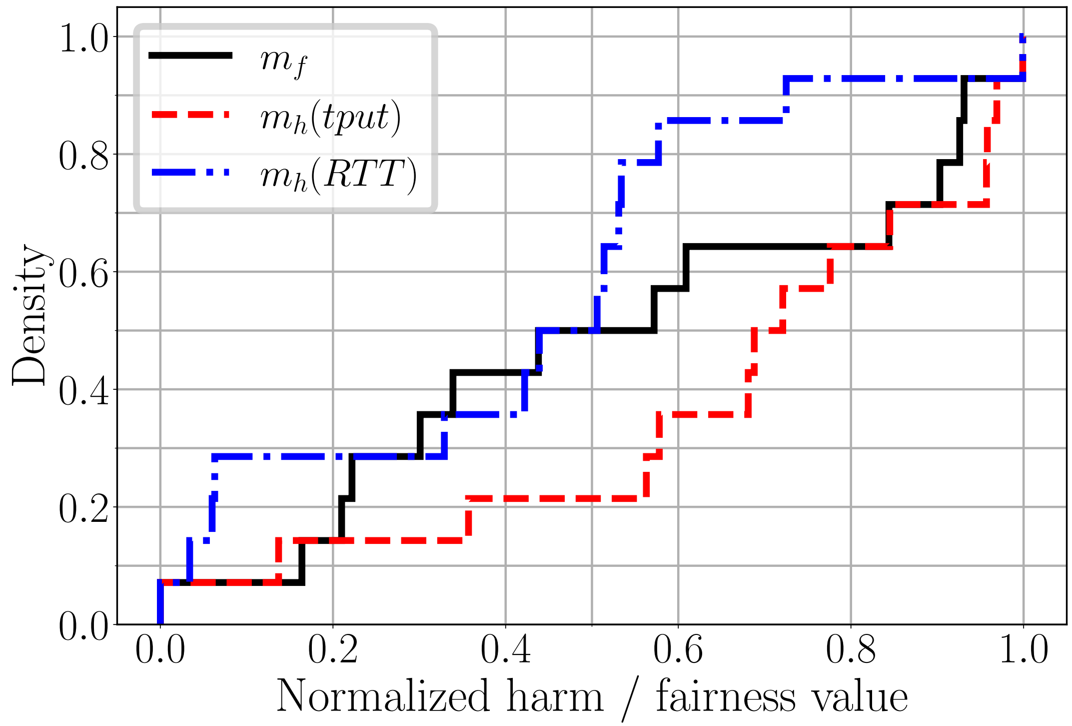
Harm and fairness distribution: cumulative distribution function of $m_h(tput)$ values for α=Cubic vs. β=Reno, α=BBR vs. β=Cubic and α=Vegas vs. β=Reno



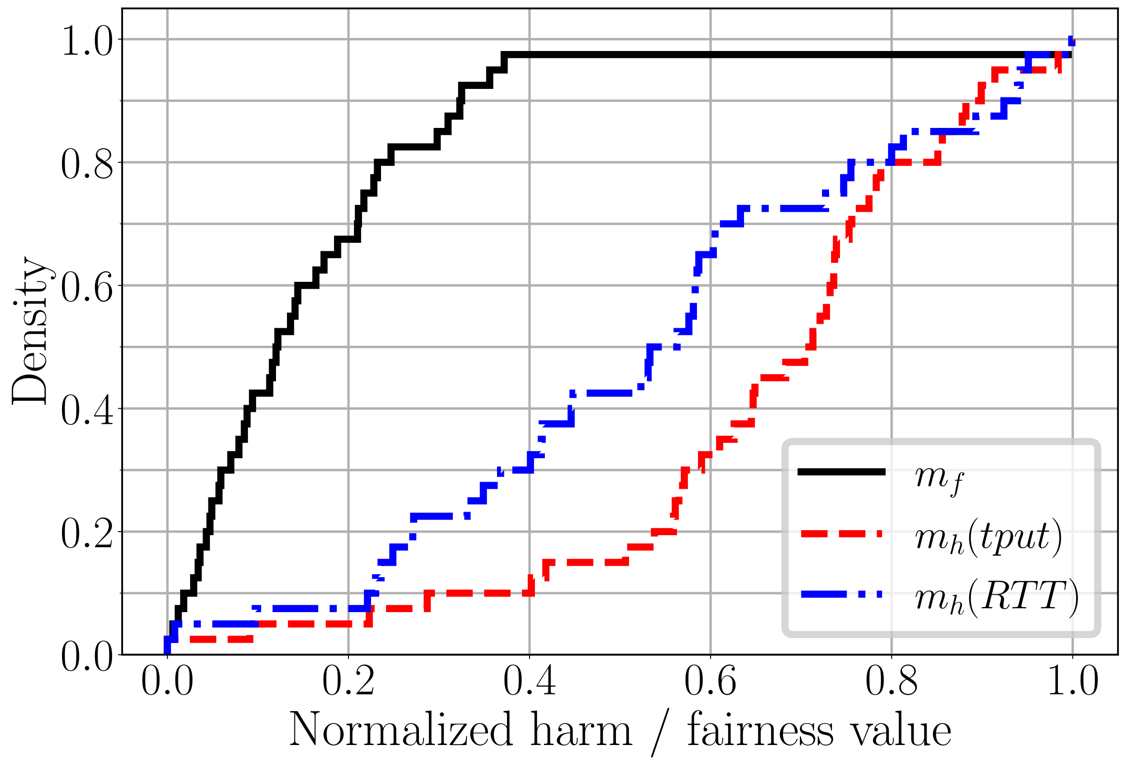
Relative harm and fairness distribution: Cumulative distribution function of normalized m_f and m_h values measured for varied α=Cubic and β=Reno pairs, across the high BDP parameter space. Larger values are better for the β flow.



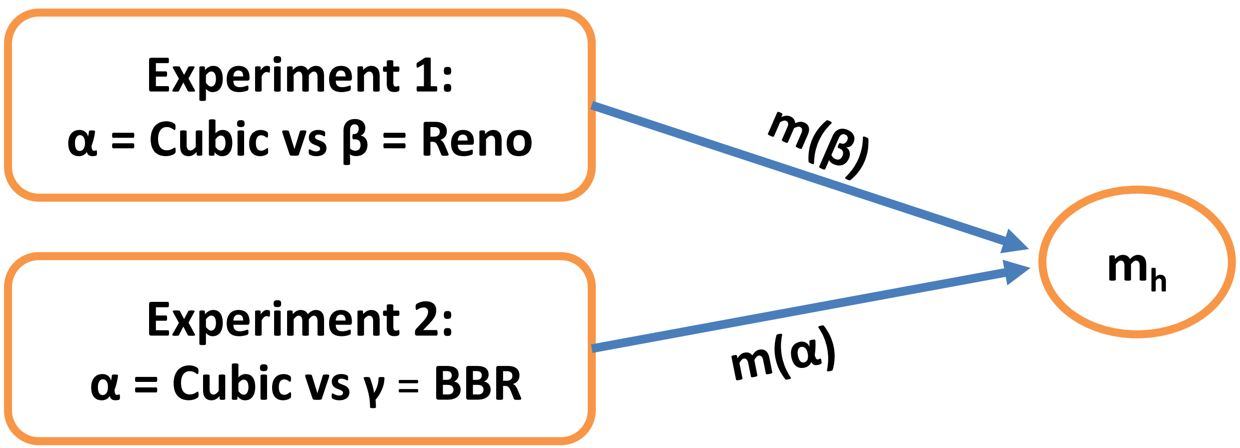
Relative harm and fairness distribution: cumulative distribution function of normalized m_f and m_h values measured for varied α=BBR and β=Cubic pairs, across the high BDP parameter space. Larger values are better for the β flow.



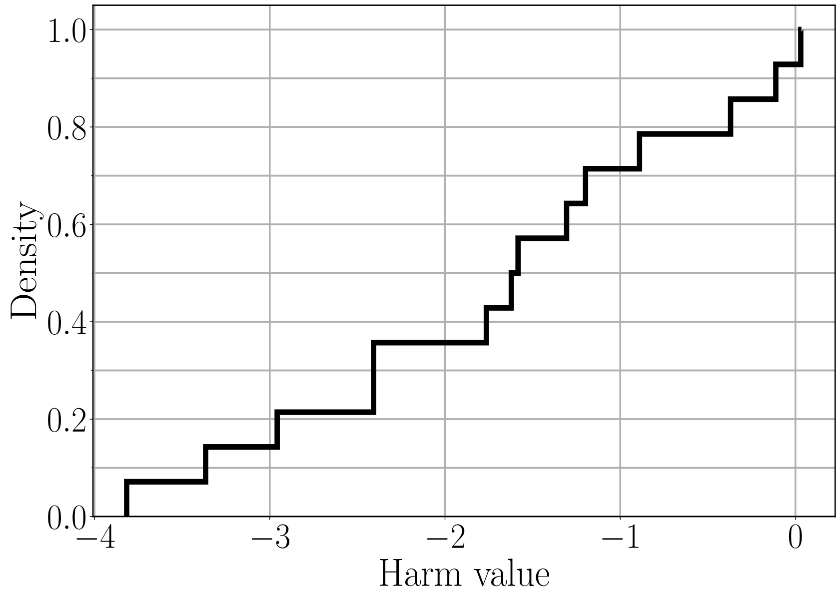
Relative harm and fairness distribution: cumulative distribution function of normalized m_f and m_h values measured for a Vegas α competing with a Reno β, across the entire parameter space. Larger values are better for the β flow.



Case study of absolute fairness and harm: scenarios that are relevant with regards to deployment of new CC algorithms



Case study of absolute fairness and harm: throughput harm comparison between (α: Cubic, β: Reno) and (α: Cubic, γ: BBR) cases in the high BDP scenario



We show raw m_h values. BBR captures more resources from Cubic than Cubic captures from Reno in nearly all cases. More specifically, it shows that BBR captures at least 1.6 times more resources for 50% of the cases and at least 2 times more in 38% of the cases.

Summary

- The harm-based approach is more useful to assess whether a next generation congestion control mechanism is safely deployable
- Presented a new linear representation of harm to better assess the differences in harm between a variety of situations
- Applied the harm concept to data produced from experiments with competing pairs of various TCP variants
  - Covered various level of aggression as well as different feedback types
  - Results show that BBR is on average 1.6 times more harmful to Cubic in high-BDP situations than Cubic is to Reno
- Plan to investigate the efficacy of harm using other performance metric, e.g., loss

Thank you!